Interpreting Coefficient Estimates in a Linear Model

Ian He

Amateur Explorer of $\mathcal{E}\text{con}\phi\text{metric}\$$

June 25, 2023

Table of Contests

Overview

Log Transformed or Not		Interpretation
Dep. var.	Indep. var.	
✗	✗	X changes by Δ units $\Rightarrow$ Y changes by $(\beta \times \Delta)$ units
✓	✗	X changes by Δ units $\Rightarrow$ Y changes by $(e^{\beta\Delta} - 1) \times 100\%$
✗	✓	X changes by $\Delta\%$ $\Rightarrow$ Y changes by $\beta \ln(1 + \Delta\%)$ units or $\Rightarrow$ Y changes by about $(\beta \times \Delta\%)$ units
✓	✓	X changes by $\Delta\%$ $\Rightarrow$ Y changes by $\left[(1 + \Delta\%)^{\beta} - 1\right] \times 100\%$ or $\Rightarrow$ Y changes by about $(\beta \times \Delta\%) \times 100\%$

Baseline Model

In the baseline model

$$Y_i = \beta_0 + \beta_1 X_{1i} + \cdots + \beta_k X_{ki} + e_i$$

the interpretation of β_1 is:

For a 1-unit increase in X_1, the expected value of Y increases by β_1 units, holding all other variables at any fixed values.

Similarly, for a 10-unit increase in X_1, the expected value of Y increases by $10\beta_1$ units, holding all other variables at any fixed values.

Only Dependent Variable Is Log Transformed

In the model

$$\ln(Y_i) = \beta_0 + \beta_1 X_{1i} + \cdots + \beta_k X_{ki} + e_i$$

the interpretation of β_1 is:

For a 1-unit increase in X_1, the expected value of Y increases by $\left(e^{\beta_1} - 1\right) \times 100\%$, holding all other variables at any fixed values.

Similarly, for a 10-unit increase in X_1, the expected value of Y increases by $\left(e^{10\beta_1} - 1\right) \times 100\%$, holding all other variables at any fixed values.

Only an Independent Variable Is Log Transformed I

In the model

$$Y_i = \beta_0 + \beta_1 \ln(X_{1i}) + \cdots + \beta_k X_{ki} + e_i$$

the interpretation of β_1 is:

For a 1% increase in X_1, the expected value of Y increases by $\beta_1 \ln(1 + 1\%)$ units, holding all other variables at any fixed values.

Similarly, for a 10% increase in X_1, the expected value of Y increases by $\beta_1 \ln(1 + 10\%)$ units, holding all other variables at any fixed values.

Only an Independent Variable Is Log Transformed II

By Taylor expansion, as $\Delta \to 0$, we have

$$\ln(1+\Delta) \approx \Delta$$

Thus, the interpretation of β_1 can also be:

For a 1% increase in X_1, the expected value of Y increases by approximately $(\beta_1 \times 1\%)$ units, holding all other variables at any fixed values.

Then, could we use this approximation in interpreting the result of a 10% increase in X_1? This is equivalent to asking whether 10% is a small value.

Both Dependent and Independent Variables Are Log Transformed I

In the model

$$\ln(Y_i) = \beta_0 + \beta_1 \ln(X_{1i}) + \cdots + \beta_k X_{ki} + e_i$$

the interpretation of β_1 is:

For a 1% increase in X_1, the expected value of Y increases by $\left[(1+1\%)^{\beta_1} - 1\right] \times 100\%$, holding all other variables at any fixed values.

Similarly, for a 10% increase in X_1, the expected value of Y increases by $\left[(1+10\%)^{\beta_1} - 1\right] \times 100\%$, holding all other variables at any fixed values.

Both Dependent and Independent Variables Are Log Transformed II

Recall an approximation method:

$$(1+\Delta)^{\beta} \approx 1+\beta\Delta$$

for a small value of $|\beta|\Delta$. Thus, the interpretation of β_1 can also be:

For a 1% increase in X_1, the expected value of Y increases by approximately $(\beta_1 \times 1\%) \times 100\%$, holding all other variables at any fixed values.

Mathematical Proof I

Suppose that our regression is $\ln(Y) = \beta_0 + \beta_1 X + e$. Let X change by Δ units (from x_1 to x_2) and Y change from y_1 to y_2.

$$\ln(y_2) - \ln(y_1) = (\beta_0 + \beta_1 x_2) - (\beta_0 + \beta_1 x_1) = \beta_1 \Delta$$

$$\Rightarrow \ln\left(\frac{y_2}{y_1}\right) = \beta_1 \Delta$$

$$\Rightarrow \frac{y_2}{y_1} = e^{\beta_1 \Delta}$$

$$\Rightarrow \frac{y_2 - y_1}{y_1} = e^{\beta_1 \Delta} - 1$$

Mathematical Proof II

Suppose that our regression is $Y = \beta_0 + \beta_1 \ln(X) + e$. Let X by $\Delta\%$ (from x_1 to x_2) and Y change from y_1 to y_2.

$$
\begin{aligned}
y_2 - y_1 &= [\beta_0 + \beta_1 \ln(x_2)] - [\beta_0 + \beta_1 \ln(x_1)] \\
&= \beta_1 [\ln(x_2) - \ln(x_1)] \\
&= \beta_1 \ln\left(\frac{x_2}{x_1}\right) \\
&= \beta_1 \ln(1 + \Delta\%) \\
&\approx \beta_1 \times \Delta\%
\end{aligned}
$$

The approximation in the last line holds when $\Delta\%$ is very small.

Mathematical Proof III

Suppose that our regression is $\ln(Y) = \beta_0 + \beta_1 \ln(X) + e$. Let X by $\Delta\%$ (from x_1 to x_2) and Y change from y_1 to y_2.

$$
\begin{aligned}
& \ln(y_2) - \ln(y_1) = \beta_1[\ln(x_2) - \ln(x_1)] \\
\Rightarrow\ & \ln\left(\frac{y_2}{y_1}\right) = \beta_1 \ln\left(\frac{x_2}{x_1}\right) \\
\Rightarrow\ & \frac{y_2}{y_1} = \left(\frac{x_2}{x_1}\right)^{\beta_1} = (1 + \Delta\%)^{\beta_1} \\
\Rightarrow\ & \frac{y_2 - y_1}{y_1} = (1 + \Delta\%)^{\beta_1} - 1 \approx \beta_1 \times \Delta\%
\end{aligned}
$$

The approximation in the last line holds when $|\beta_1|\Delta\%$ is very small.

Some Good Examples and Guidelines

- How Do I Interpret a regression model when some variables are log transformed? (UCLA: Statistical Consulting Group)
- How can I interpret log transformed variables in terms of percent change in linear regression? (UCLA: Statistical Consulting Group)
- *Interpreting Log Transformations in a Linear Model* (Clay Ford, 2018)